Knowledge vs True Belief in the Socratic Psychology of Action

Terry Penner

At *Protagoras* 351b-357e, Socrates argues that knowledge is something 'strong' which 'cannot be overcome by pleasure' (352b-c). The traditional interpretation of this passage understands this 'strength of knowledge' claim to follow from the claim that 'no one errs willingly [at getting what is best for them]',[1] or, more exactly

> ISBA No one ever acts contrary to what they, at the moment of action, believe to be the best (most advantageous) option open to them.

[I call (ISBA) the impossibility of synchronic belief-akrasia — 'akrasia' because in akrasia one is acting contrary to what one knows or believes to be best; 'belief-akrasia' because it is belief, rather than knowledge, that one is acting against; and 'synchronic' because it is what one believes *at*

---

1 That is, no one errs willingly at the science of getting what is most advantageous for one — that science which is identical with virtue. On the other hand, we *can*, and do, err willingly with respect to other sciences because we can have a motive to err willingly with respect to the goals of other sciences, but not with respect to getting what is best. See *Lesser Hippias* 367a8-b1 (all line numbers from Oxford Classical Texts) as well as my article (1973b, 138-43, and 146), on economic man not erring willingly with respect to his own economic interest. Economic man would certainly err willingly with respect to a chess game if he thought it in his economic interest to lose the chess game.

the moment of action that one is acting against. When, in modern times, philosophers speak of the alleged 'impossibility of akrasia' or the alleged 'impossibility of clear-eyed akrasia', it is in fact (ISBA) which they have in mind.] On this purely economic view of the relation of belief to action, one *automatically* acts in accordance with what one believes best at the moment of action. But then since what one knows one also believes, we get as an immediate corollary the impossibility of synchronic knowledge-akrasia:

> ISKA   No one ever acts contrary to what they, at the moment of action, know to be the best (most advantageous) option open to them.

Strange notion of strength: for what one *automatically* does! Be that as it may, it is here, in the impossibility of synchronic knowledge-akrasia, the modern interpreters have located the Socratic 'strength of knowledge'.

Now there is no question that Socrates is committed to both (ISBA) and (ISKA), since both follow trivially from 'No one errs willingly'. But is (ISKA) what Socrates has in mind when he says that knowledge is something strong? A new interpretation of the *Protagoras* suggests not.[2] Instead of this purely economic view of the strength of knowledge (and indeed the strength of belief!), the new interpretation suggests rather a notion of *epistemological* strength — the ability to hold onto one's intellectual grasp on the situation as different gestalts on the same situation successively present themselves to the agent throughout the temporal context of the action. The following example may get across the idea here. I see the oatmeal and raisin cookie on the counter as I am about to pass through the kitchen. I form the belief that it would be best for me over all *not* to have that cookie. This belief is in full force as I begin to pass through the kitchen, but just as I come close enough to see more clearly the delectable and mouthwatering features of the cookie (and, as chance — or nature — has it, I am at that instant also within closer reach of the cookie), I suddenly change my mind, reach out for it, and start eating it — saying to myself immediately it is in my mouth, 'What a fool I am'. In this sort of case,

---

2   For detailed discussion of the traditional interpretation and the new interpretation, see my (forthcoming, 1996); and for a briefer sketch, my (1992).

I do, at the moment of action, believe that it will be best to reach out, grab, and eat the cookie. My vice *at that moment* is ignorance. I do not err willingly. At the same time (so to speak), I am *not*, in this course of action, acting in accordance with what I, for most of the temporal context of the action, believe to be best for me. (The 'temporal context of the action' here includes not just the moment of action itself, but also the time immediately preceding the action, during which I contemplate the action and various options to it, and the time immediately following the action, when I may be pleased with, or regret, what I have done.) The belief I held through most of the temporal context of action proved unstable and, at the moment of action, wavered.

This little episode exemplifies the idea attributed to Socrates by the new interpretation: the idea that *mere belief* as to what is best will characteristically be unstable as differing gestalts on the cookie successively present themselves, while *knowledge* as to what is best will *not* be thus unstable. Thus according to the new interpretation, Socrates thinks we have the *possibility* of diachronic belief-akrasia:

> PDBA One *can* (voluntarily) act contrary to what one, throughout most of the temporal context of the action, believes to be the best option open to one — because of the possibility of a brief and sudden mind-change at the moment of action, the mind-change being brought on because one's beliefs as to the relative desirabilities of various options vacillate as those options are viewed from different perspectives;

and the *impossibility* of diachronic knowledge-akrasia:

> IDKA No one ever acts contrary to what they, throughout most of the temporal context of the action, *know* to be the best (most advantageous) option open to them.

Knowledge, under assault from different ways in which the contemplated action may successively present itself throughout the temporal context of the action, is stable and unwavering. Mere belief, on the other hand — even mere *true* belief — may be expected to vacillate under such assault. This new interpretation urges that it is here that we find the significance of the Socratic claim that 'the power of appearance makes us wander and often change up and down with respect to the same things and change our minds in our actions and choices' while knowledge (the 'measuring-art') 'renders the appearance ineffective and makes our soul *remain* in the truth and saves our life' (356d-e).

('Remaining', like 'stability', *takes time*: there is no such thing as stability *at the moment of action*: hence the 'temporal context of the action'.) The analogy is to judgments of size at different distances from the things to be measured, where one may be taken in, in one's judgment of relative sizes by the effects of perspective — but where one with a full knowledge of the science of perspective will not be taken in, no matter what the point of view from which the magnitudes are presented.

Now, it can be said on behalf of this view that there will be some cases of belief which are cases of weak conviction; and perhaps some of those *will* be unstable by this very fact. But what about stubborn beliefs, or fanatical beliefs? Have we any reason to think they will be less stable than the beliefs of those with knowledge? This is exactly the point Aristotle urged against those who took Socrates to hold that it is *knowledge* that is strong, while *true belief* is intrinsically weak. Unfortunately, we have no direct evidence as to how Socrates would have replied to any such objection.

It is the aim of the present paper to fill this gap — by constructing, on Socratic principles, just such a reply. Indeed, I shall suggest that, on certain not implausible assumptions, that reply will not only motivate the claim that knowledge is stable while even stubborn and fanatical belief are inherently unstable, but will actually show Socrates right! The question is thus of more than just historical interest.

I shall proceed as follows. In §I, I begin with Aristotle's version of the objection just noted in a little more detail, and sketch my strategy for responding to the objection. Then, in §§II-IV, I lay out an argument by cases, by using which, I claim, Socrates could have argued — correctly, on assumptions I explain — that knowledge is indeed stable and mere belief, even true belief, and even stubborn or fanatical true belief, is inherently unstable.

I   The Aristotelian Challenge — Why should we suppose that Firmly Held True Belief is any less Stable than Knowledge? — and some Assumptions to be Used in Attempting to Meet the Challenge

It might seem that there is no chance that Socrates is right on the point that knowledge is stable under the assaults of different misleading alternative gestalts on the same practical situation one began with, while even stubborn or fanatical true belief is inherently unstable under those assaults. Certainly Aristotle thought there was no chance that

anyone who held that knowledge was strong and true belief weak could be right.[3] For he says (*Nicomachean Ethics* VII 3, 1146b24-30):

> As for the suggestion that it is true opinion and not knowledge against which we act incontinently, that makes no difference to the argument; for some people when in a state of opinion do not hesitate, but think they know exactly. If, then, the notion is that owing to their weak conviction [διὰ τὸ ἠρέμα πιστεύειν] those who have opinion are more likely to act against their judgment than those who know, we answer that there need be no difference between knowledge and opinion in this respect; for some men are no less convinced of [πιστεύουσιν οὐδὲν ἧττον] what they think than others of what they know ...

Now it seems plain that Aristotle is thinking here of *synchronic* strength of knowledge and true belief. (It apparently does not occur to Aristotle that the knowledge-akrasia and true-belief-akrasia which Socrates is speaking of might be *diachronic*, rather than synchronic.)[4] But Aristotle's objection can be readily extended to diachronic cases. If someone's true belief is *stubborn* or *fanatical* at an instant of time, why shouldn't it be such throughout the period of deliberation?

So Aristotle's objection remains formidable when extended to cover diachronic cases. Let us call Aristotle's objection, thus extended, 'the Aristotelian Challenge':

(AC) Have we any reason to believe that knowledge is going to be any more stable throughout the temporal context of the action than stubborn or fanatical true belief, when under assault by various misleading gestalts on the situation before one?

If Socrates' position on the stability of knowledge and the instability of belief is as philosophically weak as Aristotle evidently takes such a position to be, then the contrast between knowledge and belief which I have insisted is part of the argument for the strength of knowledge must come to seem of very little philosophical interest. As a matter of

---

3 For the suggestion that Aristotle didn't even think that Socrates could even have *held* the position that knowledge is strong and true belief weak, see the references in the next note.

4 See (forthcoming, nn 7, 29).

sheer scholarship, we may have to make reference to stability in talking
about *Protagoras* 351b-7e; but what is there to be learned *philosophically*
from such considerations? If the traditional interpretation is wrong
as a matter of scholarship, does it not nevertheless capture every-
thing of *philosophical* interest in Socrates' discussion? (Socrates' really
interesting error in the passage, philosophers have usually supposed,
is his claim that [synchronic] belief-akrasia is impossible[5] — though

---

5   See, for example, the wonderful passage in Vlastos (1956, xliv-xlv) where Vlastos
    throws up as obvious counter-examples to
       (1) No one ever acts contrary to what they know *or believe* to be the best option
           open to them,
    the cases of the Euripidean heroines Medea and Phaedra, who both say that they
    are acting contrary to what they think best:
         So here on the stage before our eyes Euripides gives Socrates the lie. Who is right?
         The trouble with Socrates is not so much that he was wrong on this point (and
         I, for one, unquestionably think he was) as that his method did not provide him
         with the means by which he would be likely to correct, or, at least, suspect his
         own error. He was too fascinated by the patterns into which he could organize
         his propositions to reflect with the needed sensitiveness and humility on matters
         which can only be learned from the facts themselves or from those whose vision
         of the facts if more subtle and penetrating than one's own. Had his method been
         less narrow he might have sensed how false was his metaphor at the end of the
         mock-exegesis of Simonides, how much more than entertainment a moral
         philosopher could get from poets and others who are no great arguers but know
         the human heart.
    (See also the more extended discussion of Medea and Phaedra in Walsh (1963, 16-23,
    27) [= Vlastos (1971, 250-7, 262-3)].) Again, in his (1958), evidently still endorsed in
    1971 [Vlastos (1971, 15-16)], Vlastos puts the case just as strongly against (ISBA),
    and so against 'No one errs willingly':
         Aunt Rosie is afraid of mice, but she knows quite well that a mouse can do her
         no great harm. She knows she runs a far graver risk to life and limb when she
         drives her car down Main Street, but she is not a bit afraid of that, while she is
         terrified of mice. This is absurd, but it happens; and her knowing that it is absurd
         does not prevent it from happening either, but only adds shame and guilt to fear.
         This is not evidence of a high order; it is just a *fact* that does not square with
         Socrates' theory.
    For all their persuasive power, these two passages both rely on a notion of *plain and
    obvious empirical facts about desires* which must necessarily take it as just obvious *what
    the nature of human desire is*. True, Plato, Euripides, and Vlastos may all think it
    obvious that there are brute irrational desires and that *that* is how akrasia occurs.
    On the other hand, Davidson insists that there are only rational desires involved in
    akrasia, albeit proceeding from different partitions of the mind [(1986), (1970); also
    my (1971, §1) — in my (1992), I reject both views: cf esp. 48 with n 13 on 'the facts'.]

I have myself argued elsewhere that Socrates is *not* in error on this point.[6]) Why not say that Socrates holds that the impossibility of synchronic knowledge-akrasia is a consequence of the impossibility of synchronic belief-akrasia, and then mention, in passing, that Socrates made this other silly (and doubtfully relevant) error about stability? This error, it can then be added, is not of any great

———————

6   The view in (ISBA), that people always act in accordance with what they think best at the moment of action, has been rejected by most philosophers, from the time of Plato's parts of the soul argument [at *Republic* 436-40] down to the present. Plato and Aristotle say that besides rational desires (desires for what one thinks best, on the usual view), one can also have brute irrational desires that simply outmuscle one's rational desires. For Plato, as for Aristotle, there is simply *no problem* about how akrasia can take place, once one sees that there are irrational desires. But in spite of the facts that
   (i) we have no defence from Socrates himself against Plato's account of irrational desires;
and that
   (ii) we may reasonably suppose that Plato thought there *was* no response to this attack,
I have nevertheless argued, in my (1990, principally at 49-61, esp. 59-61), that there is a response. The basic idea of that response is that if a purely irrational desire for food, drink, or sex (of the sort Plato and Aristotle believe in) is to lead one to some particular voluntary action — and all voluntary actions *are* quite particular voluntary actions — it will have to account for how the brute desire for drink, say, can become a desire to move in *just this* direction to pick up *just this* particular drink. I suggest, first, that it is not enough merely to cite the co-occurrence of this brute desire with a particular belief, since, on any ordinary view, we have many different desires and a great many different beliefs. *Something* must decide that it is just this belief and just this desire that 'get it together' in action. I suggest that what is needed is some sort of substitution device like 'whatever is best' in this practical syllogism:
   (a) I want whatever is best in this particular situation;
   (b) What is best in this particular situation is to satisfy my thirst with *just this drink*;
So (c) I want (and pick up) *just this drink*.
Without this sort of substitution, the brute desire can't push one in the direction of this particular drink. With it, the desire for this drink which drives this action is not a brute irrational thirst for drink *simpliciter* of the sort Plato requires in his 'Parts of the Soul' argument [and to which Aristotle is also committed: see my (1990, 49-50, esp. n 14)]. Rather, what we have is a transformation of the original desire for what is best by way of the substitution into 'desires what is best' of particular beliefs about what is best in the situation. But then the desire that drives the action is this (rational) desire for good, and not the brutely irrational desire for drink *simpliciter*, which Plato and Aristotle tell us is what is operative. I shall have more to say on this topic elsewhere.

significance in Socrates' theory of action, and doesn't affect much the valuable part of Socrates' discussion (on synchronic belief-akrasia).[7] Such, I think, has been the reaction of those proponents of the Traditional Interpretation who have nevertheless, as a matter of scholarship, recognized Socrates' apparent insistence on the question of stability.

So the all but conclusive exegetical case for the new interpretation, and against the Traditional Interpretation is, in a certain way, not secure. It is not secure *philosophically*. Something must be said in defence of Socrates' (to Aristotle, quite implausible) view that even firmly held mere true belief is inherently unstable. Unfortunately, it must be granted that Socrates nowhere gives any argument *even for the stability of knowledge*. This is true, even at *Meno* 97c-98a or at *Euthyphro* 15b-c, where one might think it most important for Socrates' theory of philosophical dialectic to make out such stability. (More on the *Meno* on stability below.) So it might be thought that the notion of the stability of knowledge was relatively marginal to Socrates' ethical thought. I shall nevertheless argue that it is *central* to his ethical thought.

Beginning in the next section, I shall go through four sorts of cases where one might think one could have stable belief or unstable knowledge, and try to show that none will serve to refute Socrates' position. I begin with the sorts of cases of firmly held true belief (Cases 1 and 2 below), where what makes the true belief in question stable is that the believer *thinks* he or she *knows* — the kinds of case which, one might well think, simply play into Socrates' hands. (Some would say Socrates likes nothing better than refuting claims of those who think they know, though they do not.) The treatment of these cases generalizes in turn to Case 3, where one has a firmly held true belief *without* thinking that one knows, but thinking that one has good reasons. Case 4 is a much harder sort of case than any of the first three. Here what accounts for the *firmness* of belief is nothing *intellectual*, such as thinking one knows or has good reasons, but something having to do with *will*: stubbornness, determination, strength of will.

My discussion of Case 1 will be largely an elaboration of a particular example; so I will need the assumptions that I make in that elaboration

---

7 For another view, see Taylor (1991 [1976], 191-2), and for my response, (forthcoming, nn 23-4).

to hold generally if I am to make my case. Case 2 will also depend almost entirely on two assumptions which I shall state. And it will be clear that Case 4 and the Appendix also require some large philosophical assumptions. At the same time, none of the assumptions I shall make seem to me entirely unreasonable — either philosophically or as something Socrates would embrace. The most crucial assumption, however, is my repudiation of the existence of brute irrational desires (for food, drink, sex, warmth, and so forth) of such a sort that they can bring us voluntarily to act contrary to what we think is best for us in the situation. In n. 6 above, I have sketched the defence I give elsewhere of this assumption.

A word about my overall strategy in succeeding sections. What I shall do is exploit the systematic interconnectedness, within that knowledge (or science) of goods and bads which Socrates identifies with virtue, of *all* one's beliefs about the good and the bad. It is this systematic interconnection — which I have elsewhere, in my (1992a) and (1973b), attributed to Socrates — which, I argue, ensures that false beliefs anywhere else within this part of one's belief-structure tend to lead to the defeat of any true belief with which one may start. My idea is to argue that the difference in reasons between those with knowledge of goods and bads and those with mere true beliefs about goods and bads is that those with knowledge (the measuring art, the science of goods and bads) will have *no* relevant falsehoods about goods and bads in their belief-structure, while those with mere true beliefs will be operating with false beliefs elsewhere in their belief-structure. And these false beliefs will render the *true* beliefs in question unstable. A mistake anywhere in the believer's belief-structure can become — either through a mischievous use of something like Socratic dialectic, or merely through circumstance rendering the false belief operative — a failure at the original place where (it appeared) there was a true belief.

But let us see how this overall strategy works out with respect to the different kinds of cases it will have to deal with.

II   Cases in which those with a True Belief about What is Best Think they Know, when they do not: Case 1 (One has a mere True Belief about What is Best, but some of one's other Beliefs about what is Best are False)

As I have said, cases in which those with true belief *think* they know, when they in fact do not know, seem almost to cry out for examination by students of the philosophy of Socrates. Not only is thinking one knows, when one doesn't, a familiar phenomenon to Socrates. So is thinking one knows when one doesn't, though one does have (what we call) a 'true belief'. Indeed, I believe this is one of Socrates' favorite cases of thinking one knows when one doesn't. Thus, as I have pointed out elsewhere (1992a, 141-2 with nn 68-9, 41), Nicias has an important part of the truth in the *Laches*. It is actually true, *as far as the sentence he uses goes*, that 'Courage *is* the knowledge of the fearful and the hopeful' (*Laches* 194c8-d2 with 194d3 cf *Protagoras* 358d5-7, 360c6-d5). So it comes naturally to us to say that Nicias 'truly believes' that courage is knowledge of the fearful and the hopeful.[8] But Socrates' refutation of Nicias is surely intended by Socrates to show that, for all that, Nicias doesn't *know* that courage is the knowledge of the fearful and the hopeful. In my own account of the passage (1973a, 1992b), this is because Nicias has a false belief elsewhere in his belief-structure — falsely believing that courage is only a part of virtue. And I add here that even if Nicias were to 'acknowledge' that courage is the whole of virtue, he still might fail to know that courage is the knowledge of the fearful and the hopeful, if he can be tripped up from somewhere else in his belief structure — if, for example, he doesn't know that knowledge is something strong which cannot be overcome by pleasure.

Similarly, in the *Charmides*, Critias utters what are (or might easily be), by Socrates' lights, many true sentences (for example 'Temperance is knowing what one knows and doesn't know', 'Temperance is knowing oneself'). But he understands nothing of what he is saying, and changes up and down under Socratic questioning. (Compare the remark about the poets at *Apology* 22c2-3: καὶ γὰρ οὗτοι λέγουσι μὲν πολλὰ καὶ καλά,

---

8   To say that this comes naturally to us is not to say that it is true. But I shall reserve for another place expression of, and arguments for, my larger doubts on this question of the objects of mere true belief.

ἴσασιν δὲ οὐδὲν ὧν λέγουσι; and the similar remark about politicians, soothsayers, and diviners at *Meno* 99c1-5, d4-5.)[9] Once more, the explanation I suggest is that Critias has all sorts of other false beliefs elsewhere in his belief-structure — which false beliefs Socrates seems to think make his understanding even of the content of (what we call) his 'true beliefs' defective.

Socrates' method in refuting Nicias and Critias is the method he uses on all his interlocutors — going into that part of the interlocutor's belief-structure which is concerned with the science of goods and bads, and seeing how, as a whole, it adds up — invariably finding claims that won't stand together with other claims. As Socrates sees it, a failure gained anywhere is a failure at the original spot. (We can always gain from a failure elsewhere a contradiction with the original spot.) To know one thing about the human good is to know everything about the human good. (No wonder that the wisest person in Greece knows only that he knows nothing. [Penner (1992a), §V]) Because of the false beliefs elsewhere in the belief structures of Nicias and Critias, their (supposedly true) λόγοι don't 'remain' but 'go walking', like Daedalus' statues (*Euthyphro* 15b7ff).

What I am suggesting here is that the idea of the instability of firmly held 'true belief' for Case 1 is on clear display in the Socratic dialogues, at least on *theoretical* questions. Even Thrasymachus truly believes that the just is the advantageous, so that, *a fortiori*, justice *will actually be* the advantage of the stronger: *Republic* I 339b with 338c, 354a-c — the *sentence* 'Justice is the advantage of the stronger' will be true, even though Socrates refutes Thrasymachus when he puts it forward. And Meno, after a false start on what virtue is, does affirm what is surely, by Socrates' lights, a true sentence, namely, 'Virtue is the ability to get good things' and yet Socrates refutes Meno when Meno affirms it. (This is of course because of *another* belief of Meno's: that health, wealth, gold, silver, and high office are good things: 78b9-d3.[10])

(Notice that I am deliberately leaving aside for the moment cases where the agent's mere true belief about goods and bads if *not* accompanied by false beliefs about goods and bads elsewhere in the agent's belief-structure. Such cases are taken up under Case 2.)

---

9 There is some further discussion of the *Meno* passage later in this section.

10 See my (1973b, 149-50).

Now, so far, I have been speaking merely of the stability of *theoretical* instances of mere true beliefs that are not knowledge — such as those (perhaps even firmly held) true beliefs of Nicias and Critias. I now want to suggest that these considerations apply equally to lots of firmly held true beliefs about what to do in ethical *practice*. This will be especially clear in philosophers like Socrates (it applies also to Aristotle when he is thinking of *chosen* actions), who think of actions as done as means to ends of a certain hierarchical structure, and so as construed in terms of something like Aristotle's 'practical syllogism'. Let me explain briefly why I say this.

Since every action whatever is a quite particular action, if it is to be done as a means to what is best (in Socrates, as in Aristotle, happiness), it will in general be the case that the intellectual structure involved in the action will be representable in terms of what are, from my point of view, three sorts of premises: (a) premises about those things falling under the science of goods and bads (the general means to happiness), (b) premises about those things falling under other sciences (auxiliary general knowledge necessary to choosing the right kind of particular action for the general means being considered), and (c) premises involving particular things and circumstances. For example, when Helen refuses to drink this chocolate milkshake out of a true belief that it is best to do so, this will proceed (Aristotle suggests to us: VII 3, etc.) out of a practical syllogism in Helen of the following sort:

(2) It is best (maximizes happiness) in situations of kind *K* (temptation by milkshakes, ...) to do what maximizes health;

(3) Drinking milkshakes in situations of kind *K* will not maximize health;

(4) *This* situation is of kind *K*.

When these premises coalesce, Aristotle tells us, the conclusion is some particular action. In the present example, the particular action is *this* action of refusing the proffered milkshake. Now, we need not concern ourselves with the stability or instability of knowledge or true belief of premises of the sort of (3) and (4). Any counter-examples to the stability of knowledge or to the instability of true belief dealing with (3) or (4) will be irrelevant to the issue between Socrates and Aristotle, which is over the stability of knowledge of *goods and bads*, and the instability of mere true belief about *goods and bads*. It is easy to get distracted on this point, since akrasia is concerned with *actions*; so that, in general, considerations of the sort in (3) and (4) will always be involved, along with ethical considerations of the

sort we have in (2). So one might easily think that questions of ethical *practice* involve elements going beyond one's concerns in ethical *theory*. But questions of ethical practice are only *ethical* questions if they derive from one's beliefs about general means of the sort in (2).[11] Even if we grant counter-examples to the Socratic claim as applied to (3) and (4), the issue remains: is there an instability *in action* that is present in those with mere true belief about goods and bads, which is absent in those with the knowledge of goods and bads? It is to this point alone that we need to hold Socrates.[12]

With this much by way of explanation of why I think that there are no *ethical* considerations involved in ethical practice that are not involved in ethical theory, I can now urge that the same sort of instability we have found to supposed knowledge which is merely *firmly held true belief* in the theoretical discussions of the Socratic dialogues can also be expected in practical discussion and practical decision-making. For if Socrates

---

11 This may seem wrong. It may seem that besides ethical theory, there is ethical *perception* and ethical *sensitivity* — matters which (it might be thought) relate not to theory but to perception. As Aristotle puts it, ἐν τοῖς καθ' ἕκαστα καὶ ἐν τῇ αἰσθήσει ἡ κρίσις (1109b22-3, 1126b23-4). Aristotle seems to be saying here that everything *critical* to our judgments in a situation that we encounter comes from the *particulars*, rather than from the universals involved. I hope, however, that we don't have to take Aristotle to be, in idealist fashion, contrasting the aesthetic particular with the conceptual universal. I hope we can construe perception as involving the scrutiny of particulars precisely to find *what the relevant attributes (universals) are in terms of which we project from the situation to the action to be undertaken.* My hope here, in other words, is that Aristotle doesn't take reacting to particular situations as perceptual or intuitive *as opposed to* being a matter of apprehension of relevant *kinds.* For I take the ability to search particular situations for *all* the kinds that are relevant — so that each is given its due weight in calculating what is best over all from the point of view of happiness — is what is involved in ethical perception. (Just as what is involved in the educated perceptions of an engineer is the ability to search situations for all the kinds relevant to the engineering project.) In fact, I think there *is* some worry that Aristotle assigns far too much to apprehension of the particular and far too little to the apprehension of the universal in his account of practical thought. If so, I shall be anxious to correct him in the way indicated. The number of relevant attributes present in a situation calling for ethical decision are, in general, enormously many. (There is more about the complexity and multifacetedness of the situations we are generally faced with in the next para. of the main text, as well as in §§III, V, below.)

12 Questions about the stability of, for example, knowledge of another of the sciences, or knowledge of particular circumstances, will have to be explored elsewhere.

expects those true ethical beliefs of Nicias and Critias to be accompanied by false ethical beliefs elsewhere in their belief-structure, which *render their hold on their true ethical beliefs unstable*, so will he expect those with mere true ethical beliefs to have false ethical beliefs elsewhere in their belief structure which will make them unstable in *acting on* their true ethical beliefs. One way in which people can be led to *do* something contrary to their true ethical beliefs would thus be by their being exposed to a line of questioning in some interlocutor who finds a weak point (false presupposition) in the belief-structure of the agent, and then uses that to create confusion at the original point. This is what happens in Socratic dialectic concerned with questions of ethical theory (and indeed in the practical case involving Euthyphro's father). Another would be for the following sort of thing to happen: In a certain multi-faceted situation — and which situation we ever confront fails to be multi-faceted? — the salient aspects of the situation speak to a firmly held true belief of the agent. But then, perhaps unexpectedly, a different aspect of the situation flashes itself at the agent, thereby focussing the agent's attention on matters that bring into play a quite different, and false, belief, within the agent's belief-structure. Then, as different gestalts flash themselves at the agent, the agent begins to 'change up and down'.[13]

Let me try to illustrate the application of the point about the Socratic dialogues to practical situations, and the way in which different aspects of a situation can activate false beliefs elsewhere in one's belief-structure. Consider the following example derived from Hume (*Treatise* III xi 7, 536-7 [Nidditch]). This is in fact only a slightly elaborated version of the sort of thing Socrates is talking about in the *Protagoras*' discussion of the measuring-art. Looked at from the standpoint of January 1, 1996, the appearance to Helen of the preferability of a milkshake on January 2, 1997 over an ice cream cone on January 1, 1997, is a true one. (Helen we may imagine, has

---

13 First example: Because of some arithmetical misperception, I keep getting one result when I add the figures up from top to bottom, another when I add them up from bottom to top. Second example: When the jurors think about what the prosecutor says, they concentrate on aspects of the situation which make them think 'guilty'; but when they think about what the defence counsel says, they concentrate on aspects of the situation which make them think 'not guilty'. The jurors' belief-structures are thus disordered, and it is *pressures of time* that in this case produces agreement, or a vote, rather than that re-ordering of belief-structures which Socrates would insist on pursuing.

been convinced, by someone she trusts implicitly, of some Decision Theory as applied to events distant in time, even though she doesn't fully understand the principles involved.) But another false belief — that it is always best to seize the pleasure of the moment — becomes activated on January 1, 1997, at which point there is a *false* appearance of what is most pleasant. So Helen takes the ice cream cone, and then regrets this action on January 2, when she contemplates the milkshake she has passed up, and is once more convinced that the principles of Decision Theory should after all have been applied even on January 1, 1997.

The idea of the benefits of seizing the pleasure of the moment may seem too nugatory an underlying false belief: some may be tempted to say that the injunction doesn't really represent a belief, but merely a surge of irrational desire, so that we don't here have a Case 1 example. (If so, this would take us to Case 4 below, where it is a matter of strength of appetite confronting stubbornness of will.) So here is an only slightly more complicated example.

Both Barbara and Helen assent to the sentence:

> (2) It is best (maximizes happiness) in situations of kind K (temptation by milkshakes, …) to do what maximizes health;

and so we say, though with due caution (n. 8 above), that Barbara knows that (2) is true while Helen merely 'truly believes' that (2) is true. But Helen also believes that

> (2H) The reason health is the best means to happiness in situations of kind K is that in those situations eating right, exercizing, and staying healthy, slim, and trim is the best means to a good physical appearance — which is itself the best means to the social success of going out with the in-crowd and getting dates with good-looking people — which is itself the best means to happiness.

On the basis of (2H), let us suppose, Helen is a fanatical slimmer — she holds fanatically the beliefs in (2H) and so holds fanatically the true belief in (2). In a situation where neither the facts about milkshakes and slimness nor the particular facts about the situation are in question, she will hold fanatically, stubbornly, and truly, that she should not have the milkshake. By contrast with Helen's beliefs in (2H), Barbara knows that

> (2B) In situations of kind K (and indeed in almost any other situation), the social success of going out with the in-crowd,

and getting dates with good-looking people is a poor recipe for happiness.

Now, my suggestion is that it is the false beliefs in (2H) which will make Helen vulnerable to certain aspects of situations she will face — wherever there is a social success aspect — while Barbara will be insusceptible to such aspects. All we need to do is to produce cases where social success and health point in opposite directions.

Here is an example. Barbara and Helen enter an ice cream parlor for purely social reasons — they see someone they want to speak to sitting in the parlor. They are both offered milkshakes, and both refuse out of true beliefs in each of them corresponding to (2). Suddenly Helen notices something new about the situation — the presence of an unusually attractive ice cream salesperson, Alexander. It may now occur to Helen that she has fallen upon an extraordinary opportunity. Getting the attention of Alexander will be the beginning of relations with just the sort of beautiful person that will best enhance her happiness. That is, Helen's false belief (2H) kicks in — notice that Barbara has no corresponding belief — so that, taking a milkshake now, since it will get the attention of Alexander, provides the best opportunity for happiness. She can now bypass considerations of slimness, since the whole point of the slimness was to get just such an opportunity as she is now getting by actually *taking* the milkshake.

The situation has flashed a new gestalt at Helen, and she crumbles. She crumbles because of the presence in her belief-structure of the falsehood that social success is to be preferred to health. The presence of this falsehood allows the world to throw up a situation in which social success is achieved *against* the interest of health. So Helen's choices can then be redirected away from the health it was earlier, accidentally, directed towards.[14]

---

14 It might seem that in the case of Helen, the belief that undercuts her true belief is non-ethical — the connection between health and social success. But this false sociological belief could in fact be shared by Barbara, without its in the least affecting what Barbara does. For considerations of the connection of health with going out with the in-crowd and dates with good-looking people gets no purchase on Barbara, whose knowledge of what is worth trading for what tells her that such social success is no means to happiness. What differentiates the two cases of Helen and Barbara, then, is the false ethical beliefs in (2H).

Helen crumbles because she does not really understand the relative parts played by health, social success, and high-flying amorous adventure in the happy life. Think of Helen's belief-structure as a hierarchy of objects of desire with happiness at the top, and various general and particular means subsumed under each other depending upon particular beliefs about what those means will lead to.[15] An over-simple account might be this:

— Helen wants to not drink the milkshake as a means to staying slim;

— she wants to stay slim as a means to health and (more importantly) to general popularity and even, in extraordinary circumstances, success with the jet set; and

— she wants the latter popularity as a means to being happy.

Presupposed in this means-end structure is a series of beliefs:

— that drinking the milkshake will interfere with her staying slim;

— that staying slim will make her popular and healthy;

— that being generally popular is more important than health;

— that success with the jet set (if one is ever so lucky as to fall into that) is more important than being generally popular; and

— that such success and popularity are her best routes to maximum happiness.

---

15 I have argued elsewhere, in (1991) and in Penner & Rowe (1994), that Socrates holds (correctly) that no one ever acts on any desire other than desire for what is in fact best for one. Standard cases which we describe as cases of 'doing something out of a desire for what is only apparently best and in fact not best' have to be redescribed by Socrates as 'doing what (merely) seems best', or as 'going for something on the basis of the false belief that doing this will get what one desires, namely, what is in fact best'. In the next few paragraphs I shall not attempt such redescriptions, but merely 'speak with the vulgar' — as Socrates does, for example at passages in the *Gorgias* later than the passage I rely on for showing Socrates' real view. See for example, *Gorgias* 511b4 and 517b5 with d5. Evidently I shall need to speak more to these sorts of passages. But I reserve that to a projected account of the unity of the *Gorgias*. (I am grateful to Christopher Rowe and Jim Butler for keeping me honest here.)

It will readily be seen that some of the beliefs that Helen presupposes, even in this oversimplified hierarchical structure of means and ends, are true, some false. It may perhaps also be evident how it came about that Alexander's presence led Helen to simply toss her fanatical slimming policy. By persuading Helen that one of the above presupposed beliefs is false, we are able to substantially alter the direction of her desire. ('What, I can drink the milkshake *as the best possible means to attracting Alexander*? So in this case — extraordinarily — slimming is not the best way to achieve that for the sake of which I was undertaking the slimming, namely, social success? All right, then, I'll ignore health and slimming and drink the milkshake as a come-on!') If I want *A* as a means to *B*, then I presuppose the belief that *A* is the best means to *B*. If I learn, or become persuaded, that that belief is false and that *C* is a better means to *B*, then I lose my desire to do *B* and I undertake *C*. So anyone who can have their beliefs undermined about what is a means to what (slimming the best means to a social success that is more important than health) can have their desires redirected away from one means (slimming) to quite another (drinking the milkshake).

So Helen's false belief about social success admits of the appearance of something as a best means which will *not* lead either to slimming or to health, but only to social success. So she momentarily changes her mind about what is best, and has the milkshake. By the time she feels bloated, and realizes what she has given up, she has already seen that Alexander is a jerk, and not worth it (I omit the details), and so she is sorry she acted as she did.

Barbara, by contrast, no matter how attractive a salesperson is produced for her, puts her money on health and, within human relationships, something more than glamor. Not possessing a falsehood in her belief structure about the superiority of a life of glamor over a life of health and saner relationships, the world cannot throw up a means to social success as preferable to a means to health.

Let me repeat: It is not — at least in Case 1 — a question of stubbornness (as Aristotle seems to imply). Helen may be absolutely fanatical about slimming. But because the whole point of slimming is for things like glamor, hobnobbing with the jet-setters, amorous adventure, and so forth, it is possible for the world (and Alexander) to play on Helen's belief-structure.

These examples suggest, then, that, for this sort of case at any rate, Aristotle is wrong to frame the issue of knowledge vs true belief in terms of which is more stubborn — as if, in the Platonic manner, the issue were one of brute motivational strength. The issue is not one of strength of the

sort: *stubbornness* or *will power*; but of steadiness of vision as one is successively exposed to various aspects of the situation one is in. It is a question of the stability of one's grasp on the truth about the situation to the extent that it involves truths about happiness, pleasure, danger, friendship, courage, the relative merits of health and hobnobbing with beautiful people, and so forth. Knowledge about these kinds of things is the only sort of knowledge relevant here. Such stability is not possible if both

(a) there is falsity within the ethical part of the belief structure; and

(b) the situation is in any way multi-faceted.[16]

If both of these conditions are fulfilled, the same belief-structure built around hobnobbing with beautiful people will have us fluctuating 'up and down', depending upon which aspect of the situation brings itself to the fore.

In this example, then, we have in (2) a true belief — perhaps even a firmly held true belief, since Helen never expects the means to social success and health to diverge — which is nevertheless unstable because of the false beliefs in (2H). That is, we have the same sort of true belief, undercut by the presence of a false belief elsewhere in the belief structure [that is, in (2H)], which we noticed in Socrates' discussions with Nicias and Critias. In addition, both the true belief in question and the false belief which undercuts it fall under the purview of the science of goods and bads: so that it is the same kinds of beliefs that are in question here, in a practical situation, as are in question in Socrates' inquiries in the dialogues.

Thus, I believe there is good reason for thinking that the Socratic dialogues themselves show that, with lots of examples of *firmly held mere true beliefs* that fall under Case 1, it is Socrates who is right and Aristotle who is wrong. To be more precise, this case covers all cases

---

16 In a sufficiently simple situation, we will not be able to get different facets of the situation suggesting now one course of action, now another. Contrast examples like building the Firth of Forth bridge in §III.

of firmly held true beliefs where the believer thinks he or she *knows*, but where the believer does *not* know because of false beliefs elsewhere in the ethical part of their belief-structure.

Before leaving Case 1, it is perhaps worth our while to note briefly the claim at *Meno* 96dff that knowledge is stable while true belief (presumably all true belief, even firmly held true belief) is unstable. As I have said, I grant that Socrates gives no argument for this claim. On the other hand, it seems to me that interpreters of this passage have tended to treat instability in this passage in the same (mistakenly) glancing fashion as they have treated stability in the *Protagoras*. Most have mistakenly concluded that Socrates himself holds the view that he has sucked Meno into at 97c11-d3 — that there is no difference between knowledge and true belief *in practice*. In my view, the truth is that this view has been decisively refuted at 97e6-98b6 *on grounds of stability*: '[True beliefs] are *not worth much* until one ties them down with an account of the cause'. (Of what use will true belief be as a guide if it is always susceptible to change to the opposite position? Having made the point that true beliefs will guide us correctly *so long as they be* true beliefs [97b5, c10, e6-7], Socrates himself says true beliefs are 'not worth much' until they have been converted into knowledge, when they become μόνιμοι [98a3-6].)[17]

Of course, that interpreters get drawn into the view that Socrates thinks true belief is in practice as good as knowledge is partly the mischief of Plato's consummate art, which, after the decisive conclusion at 97e6-98b6, has Socrates immediately [98b7-9, with c1-2, c8-10, e12, 99a1-2] proceed as if true belief *is* as good a guide as knowledge. Socrates then sucks Meno into the delicious irony of locating the virtue of politicians in a true belief that comes, by 'divine dispensation' (as with soothsayers, prophets, poets, interpreters like Ion, and those who

---

17 Notice the emphasis with which Socrates speaks at 98b2-5: That true belief is a different sort of thing from knowledge isn't even a conjecture for him; if he knows anything — and there are few things that he knows — this one thing he does know.
    I take it that the 'tieing down' involved is *not*: making one hold stubbornly to the belief, as in Case 4, but rather getting one's belief-structure in order so as not to be susceptible to refutation as in Case 1. (Compare with the reference to 'recollection' at 98a4 with the talk of 'recollecting in order — the way in which one should recollect' at 82e12-13.)

are possessed, via a direct revelation from God), to those who understand nothing and don't know what they are talking about [99b-100a, cf Ion's generalship at 540dff esp. 541e-542a].[18]

III Case 2: One Thinks one Knows, but has a Mere true Belief about what is Best, yet all one's other Beliefs are also True. And Case 3: = Case 2 without Thinking one Knows

Case 2: In the examples considered under Case 1, I have followed Socrates in supposing that a person who thinks he or she knows, but actually has just a firmly held mere true belief, will have a false belief elsewhere within his or her belief-structure (or at least that part of the belief-structure having to do with universal truths about goods and bads). This falsehood then gives either one's interlocutor or the situation before one the opportunity to exploit that false belief to make the believer 'change up and down'. But what if the person with mere true belief has only *true* beliefs in the rest of the relevant part of their belief-structure? This is Case 2.

My idea here is to argue that, on certain Socratic assumptions which are not in themselves implausible, this case never occurs. (Certainly if there are occurrences of this case, they will escape the difficulties I noted above in my discussion of Nicias' true belief above.) My argument that Case 2 never occurs is built on two principal assumptions. The first is that

> (A1) With anything so complicated as the science of goods and bads (the science of happiness),[19] it would be impossible for one to have all the true beliefs of the science without having

---

18 Of course those who do not agree that Socrates' identification of true belief with 'divine dispensation' constitutes a mocking refutation of true belief — and Socrates evidently thought that Anytus *would*: 99e2 — will not find persuasive what I have said here. I have defended the view that the treatment of true belief in this passage is entirely ironic in my (1988, 316-9). Taylor, 200, would apparently disagree.

19 It will be seen that I am here assuming that the science of goods and bads is not distinct from the science of what sorts of things in a complex human life give one the best shot at happiness — the science of what Vlastos refers to (n. 5 above) as 'the human heart'. Cf also n. 1 above.

> them structured and understood in a way that could only
> happen with one who has knowledge and understanding of
> how all those beliefs fit together; so that there couldn't be two
> people, one of whom had comprehensive general knowledge
> of the goods and bads of human life, and the other of whom
> had no knowledge of any of these things, but did truly believe
> all the same things as the person with knowledge.

To see this, I have suggested in my (1988, App. III) and in (1991, n. 18),
consider the science of engineering. Imagine someone who, without any
training in engineering, just lucks into providing the right engineering
specifications for building something like the Firth of Forth Bridge. Few
will suppose that, in the real world, such a possibility will ever be
realized. But that is to say that there can be no such thing as *systematic
true belief* about engineering without the *knowledge* of engineering. My
claim, then, is that the same applies to the science of the goods and bads
of human life. Living as well as possible is no easier without knowledge
than carrying out complex engineering projects is without knowledge of
the science of engineering.

   This deals with the case of persons who have all the right true beliefs
and no knowledge. What about the possibility of knowledge of some
things with only true belief about other things? Here I would need to
emphasize the point already made at the beginning of (A1), that in a
full, detailed theory, different principles so much interlock that it's
hard to get knowledge of one of them without seeing how it works
together with the others. I have argued elsewhere, for example, in
(1973b) and (1992a), that Socratic knowledge of the good and bad has
this feature.

   What assumption (A1) tells us, then, is that if Barbara knows that it is
best not to drink this milkshake on the basis of a comprehensive general
knowledge of the good and the bad, while Helen merely truly believes
that it is best not to have this milkshake — or if Barbara knows that it is
best to do the action that will best lead to health, while Helen merely
truly believes it — then there will be at least some beliefs that Barbara
and Helen do not share.

   Given this assumption, I proceed to make a second assumption. This
is that

>   (A2) where the beliefs involved are of sufficient generality (as with
>   beliefs on what kinds of things it would be best for me to do
>   from the point of view of the rest of my life), this will mean

that there will be some beliefs where a person who knows and a person who merely truly believes disagree.

That is, it's not just that one person hasn't thought about the matter on which the other person holds a belief. In fact, there will be some things on which one person has knowledge and the other will have a false belief.

Returning to the present case, if the beliefs in question are to be relevant in the way already suggested, they will have to be beliefs as to what is good or bad in general. I can then use the beliefs Barbara and Helen disagree over to show ways in which Barbara will remain steady over not drinking the milkshake and over doing what health requires, while Helen can be made to become unsteady on these matters. For if (A1) and (A2) are in general true, then all the cases Socrates will have to face will be cases where the person with true belief which is not knowledge will co-occur with other false beliefs about goods and bads — false beliefs which can then be played upon to show that stubborn true beliefs can then be unstable.

On these assumptions, then, it will turn out that Case 2 collapses into Case 1 (as Socrates would surely suppose). Of course, if assumptions (A1) and (A2) are mistaken, then my argument for the instability of firmly held true belief will fail. Even so, it does seem to be reasonable to attribute these assumptions to Socrates.

Case 3: So far I have been dealing just with the case Aristotle brought up, of the person who, thinking he or she knows, is just as convinced as the person who knows. But it seems clear that the same considerations will apply also to those more modest believers, who just think they have good reason to believe what they believe, without claiming to know it.

IV Case 4: Where Firmness of Belief Flows Partly from Sheer Strength of Will

So far, I have given only intellectual accounts of stability and instability. But what if, as in Aristotle's account of akrasia, there is operative in the firmness with which someone holds on to a true belief, some non-rational factor — for example a stubborn or determined or strong-willed attitude to certain of one's beliefs? This would ally the problem of firmly held mere true belief with the problem of being overcome by pleasure when this is construed — as by almost everyone after Socrates, including the Plato of *Republic* IV — as a matter of the relative strengths

of rational and irrational desires (nn. 5-6 above). Like Plato, Aristotle understands being overcome by pleasure as occurring only by *appetite or anger* being contrary to that desire for good which aspires to being the agent's choice; while Socrates understands the phenomenon to be produced by a brief mind-change at the instant of action, always in accordance with the agent's desire for good (together with the agent's beliefs, at the instant of action, about what is best). So too, presumably Socrates will want to account for the phenomenon of *firmly held mere true belief* in the same sort of purely intellectualist terms; and those anxious to shore up the Aristotelian objection will correspondingly presuppose that this firmness often has some other-than-intellectualist explanation. Such philosophers might suppose that some people with firm true belief about the good are just *strong-willed*. (One thinks here, though one *need* not, of Aristotle's ἰσχυρογνώμονες; *Nicomachean Ethics* 1151b5.) Here we have a notion something like that of *will* which is entirely absent, and could have no function in, Socrates' entirely intellectualist account of desire.

Now, given that such notions of strength of will are just the obverse of the notion of weakness of will, it becomes natural to suggest that the anti-Socratic argument about strong-willed holdings of true beliefs will work if and only if the argument against Socrates on akrasia will work. On this showing, Aristotle's objection will introduce no new issues beyond those raised in n. 6.

All the same, it is reasonable to expect, from a defender of Socrates, some plausible intellectualist account of what appear to be Case 4 occurrences of firmness of true belief. What is this stubborn holding onto true beliefs which is supposed to be not just due to the person's reasons? Evidently, I want to explain this stubbornness, or will power, solely by means of appeal to the person's reasons. Let us go back to Helen, who, as we saw in (2) above, firmly and truly believes that it is best not to drink the milkshake because she truly believes that seeking health is her best means to happiness on these sorts of occasions (believing firmly, as she does, that seeking health in such situations is the best means to a good physical appearance and to social success, which last is the best means to happiness). Now, I think I know one way in which Helen can come to seem to have will power. We find we can get nowhere with her by saying, 'Come on, Helen; never mind about slimness and health and all that stuff. The milkshake will taste really great!' Phaedra, to whom we say the similar thing, *does* give in, and does drink the milkshake. So we say that Helen has will power while Phaedra does not. But I think it is evident that we don't *need* to

appeal to will power to understand this case. The reason we don't get to Helen with this line of persuasion is that our appeal to gustatory pleasure does nothing to offset her reasons for acting that have to do with social success. Saying such things to Helen will have all the usefulness of shouting the same words ever louder at a foreigner when the foreigner does not understand the words the first time. It is because of the foredoomed failure of such a plan of attack in Helen's case that one thinks that her true belief is stubborn and stable and that certain people have 'will power'. People like Phaedra, who do give in to such an approach will then be said to 'lack will power'. My suggestion is, of course, that Phaedra will have a different belief structure from Helen's — rating the goodness of gustatory pleasure (or of immediate gustatory pleasure) higher than Helen does in relation to the pleasures of social success. (Social success has only a prima facie appeal to Phaedra, by comparison with immediate gustatory pleasure.)

Of course, as we have seen, none of the above stops it being the case that Helen can be made to waver by finding *another* point in her belief-structure where one can press on a false belief and bring the whole structure down like a house of cards — as with Alexander. But finding such a point with someone may be quite difficult — and even require the production of a quite unpredictable situation. So we say she is 'strong-willed'.

This shows one sort of way of trying to handle putative cases of stubborn belief. By itself, it does not constitute a complete reply to examples falling under Case 4. One would have to carry out a similar analysis in all other purported counter-examples. Certainly, similar examples can be found even in Euripides — whom Vlastos has exhorted us to find as *the* great relevant source of examples of the possibility of akrasia-as-understood-by-moderns [n. 5 above]. Who could be more stubborn and stiff-necked than Pentheus in the *Bacchae* in his determination to rid his state of this orgiastic Asiatic religion? Here is a tyrant *filled*, one would say, with stubbornness and brute motivational strength ('will-power'). But Dionysus knows just how to play on Pentheus' faulty belief-structure — a belief-structure that excoriates women's orgies, but also sees some good in an amount of what one may suspect to be a little prurience and voyeurism, especially when it may enable one to scout out the military situation a little better; and so Pentheus ends up slinking through the backstreets dressed as a woman (and having even worse things happen to him). Again, who could be more stubborn and stiff-necked than Agammem-non in Aeschylus' play? But Clytemnestra knows just how to play on

his belief-structure too, and soon Agammemnon is walking on the purple.[20] Philosophers ransacking Greek tragedy for counter-examples to Socrates have attended too much to supposed weakness (see n. 5 above on Medea and Phaedra), and too little to the crumbling of stubbornness.

The source for the imputations of stubbornness in the above examples is failing to see the different underlying belief-structures, making the belief under attack susceptible in different ways to different attacks. Another example of such imputations can be seen emerging from the following sort of educational dialectic. Think of the way in which one may insist on a certain (mistaken) intellectual position, no matter what is said against it (firm belief) — until such point as an unnoticed presupposition is exposed and shown false. 'Oh, well, if I'm wrong about that,' one suddenly realizes, 'then the rest of what I'm saying *will* collapse.' So far, this stubborn holding onto a belief — which is nevertheless not stable — has a merely intellectual explanation. It *isn't really* a matter of will. It's a matter of not seeing one's way through the whole problem, in spite of thinking one has a grasp on it. (This is a matter of instability in a firmly held *false* belief, undercut by bringing out a truth elsewhere; but the considerations involved are presumably parallel to those where we have a firmly held true belief undercut by a false belief elsewhere.)

From a Socratic point of view, just such a diagnosis may seem to be required of teacher-student dynamics where the teacher attributes an educational failure to a student's stubbornness. What is called for, by Socratic lights, is the finding of that unnoticed false presupposition which, once exposed, causes the scales to fall from the student's eyes — not the thinking of unbecoming thoughts about student pigheadedness (as if the student were somehow a naughty child, or as if the student were being *disobedient*). Thus do we get imputations of stubbornness and willfulness in cases of holding onto a belief which *should* have a merely intellectual explanation — catering to a voluntarist account of firmness

---

20 Clytemnestra's obvious and masterly grasp of how to 'press Agammemnon's buttons', I am suggesting, is a knowledge of the weak points in Agammemnon's belief-structure. ('What a fool the man is,' she says to herself, 'to yield to my arguments' as he steps on the purple.)

of conviction, with teachers confusing their own frustration, impatience, or lack of skill with willful stubbornness in their students. Or so I believe Socrates would construe the situation.[21]

In this sort of example, however much social considerations (getting on with the course) may seem to counsel imputations of willfulness, the source of the logjam is actually failure to sufficiently articulate an underlying belief-structure. Those who, unlike Socrates, seek to base ethics, in whole or in part, upon duties, rules, moral laws, obligations, rights, obedience, and punishment will be especially tempted by imputations of willfulness. A Socratic may perhaps be forgiven for seeing in the unbecoming reactions to students which I have been describing, the same impulse Socrates would have seen in Plato when, in his political theory, Plato gives up on Socratic intellectualism and begins to insist on *control* and *obedience*, instead of understanding, in those who are ruled.

The examples I have adduced here under Case 4 hardly constitute a refutation of the possibility of an anti-Socratic, non-intellectualist account of stubbornness of belief or strength of will. As I have already indicated, I suspect that the dispute between Socrates and those who speak of a non-intellectualist strength of will is exactly parallel to the dispute between Socrates and Aristotle over weakness of will; and I have already sided with Socrates on the latter issue (n. 6). All I have tried to do here in dealing with Case 4 here is to illustrate how I would try to defend Socrates with respect to particular examples. A still fuller account is needed of how it is that people get the notion of a 'will-power' or of a stubbornness of belief as a primitive notion of their psychology. Anyone who notes Aristotle's wonderfully articulated

---

21 What about those who hear the teacher but are not listening? Or those who refuse to take part further in the dialectic? What of Thrasymachus or Callicles? Is this not genuine stubbornness? That is, are these not at least sometimes phenomena of will? (I am indebted to Christopher Rowe for inviting me to face these considerations.) Evidently a Socratic reply to these considerations will require that this apparent willfulness be located in the agent's belief-structure. But how is this to be accomplished? What leaves hope for the Socratic here is that notion that every action, including such actions as terminating conversations, is done as a means (ultimately) to happiness. So there is presumably a false belief somewhere in the agent's belief-structure about the merits of terminating the conversation; and this false belief — if the conversation *does* continue — Socrates has an opportunity to undercut. This would be to trace the stubbornness to an intellectual error.

behavioral phenomenologies of ἀκρατεῖς, ἀκόλαστοι, ἰσχυρογνώμονες, μαλακοί, ἐκστατικοί, μελαγχολικοί, προπετεῖς, and those with firm conviction (along with the opposites to all these), will be aware of the magnitude of the task, which I hope I do not underestimate. I cannot meet this challenge in the present paper.[22]

To sum up, then, the idea I attribute to Socrates is that both 'weakness of will' and 'firmness of belief' (or 'strong-willed' holding onto a belief) *do* occur; but what they are are particular phenomena which, in spite of first appearances, are not phenomena of will, but phenomena of more or less ordered belief-structures — with falsehood and truths differently located in different people, resulting in different forms and degrees of stability and instability. (Barbara, Helen, Phaedra.) The falsity of Helen's whole position on happiness in (2H) enables facets of the world to lay a snare for her — as Dionysus and Clytemnestra laid snares for Pentheus and Agammemnon, and as Socrates laid intellectual snares for Nicias and Critias. One with the knowledge of goods and bads is presumably not susceptible to these sorts of dangers. Such a person keeps a steady vision of what is worth trading for what, and neither cross-examination by Socrates nor the presentation of different aspects of the world will enable one to be thrown off as a result of what one believes about goods and bads. The strength of knowledge resides in its stability, and that stability in turn resides in its knowing what is worth trading for what, from no matter what point of view your choices are presented.[23]

This completes my response to Aristotle's objection with respect to Case 4. I suggest that, while it is hardly a conclusive argument against

---

22   I hope eventually to exploit the doubts noted briefly in n. 8 above to cast further doubt on Aristotle's belief in the possibility of stubborn or fanatical true belief being just as firmly held as is knowledge. But I cannot do so here.

23   In the *Gorgias*, Socrates seems to speak interchangeably of 'changing' one's views, or 'changing up and down' vs 'remaining', whether he is thinking about arriving at general ethical truths (491b6-c2, 499b9-c2, 527d4 with e7) or about arriving at a choice of action (481d7-e3 with 482a6-b1). [At 493a3-5, 'up and down' is referred not to (as I take it, Socratic) changes of opinion, but to irrational appetites in another part of the soul, apparently of the Platonic sort familiar from *Republic* 436-40.] The *Protagoras* apparently also allows us to notice changing 'up and down' as now practical, due to the power of appearance (our original passage: 356d5-e4), now theoretical, due to not knowing what virtue is (361c1-3). And with the 'wandering' in our original (practical) passage, compare the theoretical passage *Lesser Hippias* 376c2-6 with 372d7-e3. The 'up and down' twistings of *Ion* 541e7-8 are theoretical.

Aristotle, there is enough here to make it at least as plausible that Socrates has an answer to Aristotle on firmly held true belief as that he has an answer to him on weakness of will.

V Conclusion

I conclude, first, that there is at any rate *some* philosophical case for the position attributed to Socrates by the Stability View — the position that the science of goods and bads (which Socrates thinks virtue consists in) will indeed be stable when exposed to the presentation of misleading perspectives on a situation in which one is to act; while mere true belief about what is best in the situation, even if stubborn and fanatical, will not be stable. One key feature of my argument for that conclusion, without which my conclusion won't hold, is that

> (A3) Situations in which one is to act are in general of enormous complexity and so can present a wide range of potentially misleading perspectives.

They require the utmost intellectual surefootedness, and the utmost clarity in seeing how the different parts of one's belief-structure fit together. 'Will power' (as usually construed) is, I have argued, beside the point. Questions of will-power will all be taken by Socrates to go back to questions of belief-structure. On the Socratic view of knowledge, I have urged both here and elsewhere, to know one thing about the human good is to know everything.

Given this first conclusion, I feel able thence to conclude, second, that the picture attributed to Socrates on the Stability View is not disqualified on philosophical grounds from the position of clear superiority which I believe I have established elsewhere on exegetical grounds.

A final conclusion relates to the alleged distinction between theoretical and practical knowledge. Walsh (1963, 30-1) speaks of *Meno* 98a (on true opinion being as good in practice as knowledge, except for its instability) as being an amplification of the strength of knowledge in the *Protagoras* — except that strength in the *Meno* is epistemological rather than motivational, while in the *Protagoras*, strength is motivational rather than epistemological. From the point of view of the paragraph before last, this is exactly the wrong way to express the relation between the *Meno* and the *Protagoras*. In both passages the

strength in question is purely epistemological. But — and this is crucial — it is the same epistemological strength that enables one to hold out (a) against the blandishments of tempting but false alternative perspectives in practical situations involving pleasure, fear, and the like, and (b) on purely theoretical matters about ethics, in Socratic dialogue.[24] In Socrates there is no room for that characteristically Aristotelian distinction, between *practical reason* and *theoretical reason*. The sort of strength and stability of judgment which is involved in not being overcome by pleasure is no different from the sort of stability which Socrates' interlocutors aspire to, but never attain, in the dialogues. All of these considerations concerning discussion of ethical theory apply equally to ethical practise. When people are overcome by pleasure and vacillate — 'change up and down' — it is because someone, or some feature of the situation before them, finds the weak spot in a person's belief-structure.[25] Not 'will-power' (as usually construed), but reason, is what makes us strong in the realm of goods and bads.

Up to the present, there has been relatively little exploration in Western philosophy of the unique account Socrates gives of the place of reason in ethics, or indeed of the place of *knowledge* within ethics and the theory of motivation. I believe this is mainly due to the feeling that Socrates was just a lost cause, philosophically — that Socrates was so

---

24   Another way to test Socrates' view here, besides making sure it can give accounts of the Aristotelian cases cited in the preceding paragraph, would be to attempt to explain all the cases of akrasia and stubborn belief to be found in that great poet of akrasia, George Eliot. It is a long-range goal of mine to try to show that in her accounts of akrasia we find, almost without exception, that the akratic act has an explanation in terms of the agent being taken in *intellectually* by a facet of the (multi-faceted) situation he or she had not seen beforehand.

25   The reference to 'vacillation' in my account of weakness suggests a comparison between Spinoza and the Socrates I am presenting here — different as the theories of desire in Spinoza and Socrates may seem to be in other respects. (Here I am indebted to Dennis Stampe.) On the question of whether anything less than *knowledge* could save one from 'vacillations', Spinoza too, I think, would have been with the Socrates presented in this paper — a further suggestion of the philosophical coherence of Socrates' position.

obviously wrong: whether on the point that synchronic belief-akrasia is impossible or on the point that knowledge could be stable while mere true belief is not. I believe it is now time to start exploring this Socratic picture rather more fully.[26]

Department of Philosophy
5139 Helen C. White Hall
University of Wisconsin
Madison, WI 53706
U.S.A.
e-mail: tmpenner@facstaff.wisc.edu

Bibliography

Donald Davidson, 'How is Weakness of the Will Possible', in Davidson's *Actions & Events* (Oxford, 1980) 21-42, reprinted from Joel Feinberg, ed., *Moral Concepts* (Oxford, 1970).

—, 'Deception and Division', in Jon Elster, ed., *The Multiple Self* (Cambridge University Press, Cambridge, 1986).

Keith Lehrer and Thomas D. Paxson Jr., 'Knowledge: Undefeated Justified True Belief', *Journal of Philosophy* 66 (1968) 225-37.

Terry Penner, 'Plato on Thought and Desire', in G. Vlastos, ed., *Plato*, Vol. II, (New York, 1971).

—, 'The Unity of Virtue', *Philosophical Review* (1973) 35-68, reprinted in Hugh Benson, ed., *Essays on the Philosophy of Socrates* (Oxford University Press, 1992) [= (1973a)].

—, 'Socrates on Virtue and Motivation', in E. Lee, A. Mourelatos, and R. Rorty, eds., *Exegesis and Argument, Essays for Gregory Vlastos, Phronesis*, Supplementary Volume I, (Van Gorcum Assen, 1973) 133-51 [= (1973b)].

—, 'Socrates on the Impossibility of Belief-Relative Sciences', John J. Cleary, ed., *Boston Area Colloquium for Ancient Philosophy* (University Press of America, 1988) Volume III, 263-325.

—, 'Plato and Davidson: Parts of the Soul and Weakness of Will', in David Copp, ed., *Canadian Philosophers: Celebrating Twenty Years of the Canadian Journal of Philosophy*, Supplementary Volume 16 of the *Canadian Journal of Philosophy* (1990) 35-74.

—, 'Desire and Power in Socrates: The Argument of *Gorgias* 466a-468e that Orators and Tyrants have no Power in the City', *Apeiron* 224 (1991) 147-202.

—, 'Socrates and the Early Dialogues', in Richard Kraut, ed., *The Cambridge Companion to Plato* (Cambridge, 1991) 121-69 [= (1992a)].

—, 'What Laches and Nicias Miss; and whether Socrates thinks Courage is only a Part of Virtue', *Ancient Philosophy*, 12 (1992) 1-27 [= (1992b)].

—, 'Socrates on the Strength of Knowledge: *Protagoras* 351b-357e', (Forthcoming, *Archiv für Geschichte der Philosophie*, 1996.)

— & C.J. Rowe, 'Socrates on Desire: Is the *Meno* Consistent with the *Gorgias*?', *Phronesis* 29 (1994) 1-25.

C.C.W. Taylor (trans. & ed.), *Plato: Protagoras* (Oxford, 1976), second edition: 1991.

Gregory Vlastos, Introduction to *Plato, Protagoras* (Bobbs Merrill, Indianapolis, 1956), vii-lviii.

—, 'The Paradox of Socrates', *Queen's Quarterly* (1958), reprinted in Vlastos (1971) 1-21.

—, ed., *The Philosophy of Socrates* (New York, 1971).

James J. Walsh, *Aristotle's Conception of Moral Weakness* (New York, 1963), ch. 1, reprinted in Vlastos (1971) 235-63.